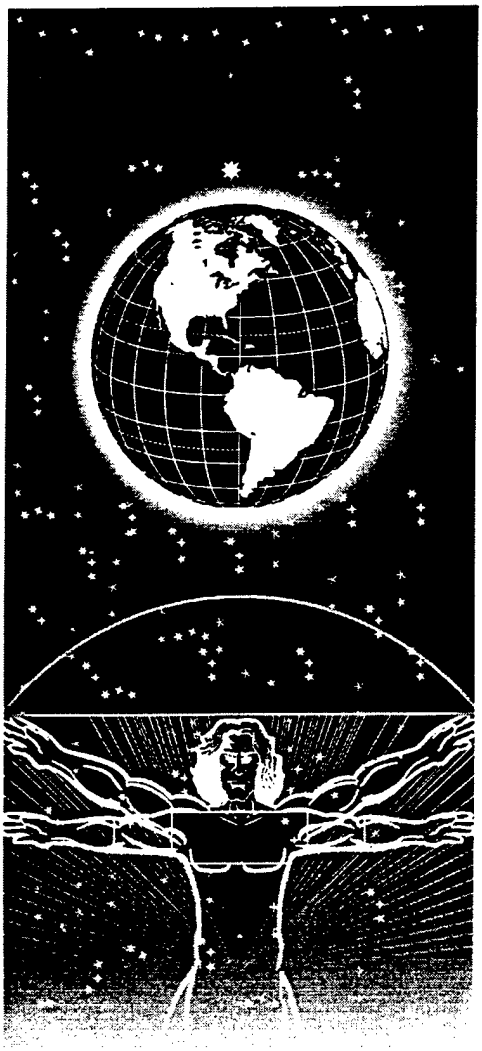
AL/HR-TP-1997-0006



UNITED STATES AIR FORCE RESEARCH LABORATORY

NEAR IDENTITY OF COGNITIVE STRUCTURE IN SEX AND ETHNIC GROUPS

Thomas R. Carretta

Training Effectiveness Branch
Warfighter Training Research Division

Malcolm J. Ree

Force Development Branch
Cognition and Performance Division

HUMAN EFFECTIVENESS DIRECTORATE
7909 Lindbergh Drive
Brooks AFB TX 78235-5352

August 1998

AIR FORCE MATERIEL COMMAND
AIR FORCE RESEARCH LABORATORY
HUMAN EFFECTIVENESS DIRECTORATE
WARFIGHTER TRAINING RESEARCH DIVISION
6001 South Power Road, Building 558
Mesa AZ 85206-0904

NOTICES

This report is published in the interest of scientific and technical information exchange and does not constitute approval or disapproval of its ideas or findings.

Using Government drawings, specifications, or other data included in this document for any purpose other than Government-related procurement does not in any way obligate the US Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation, or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

THOMAS R. CARRETTA
Project Scientist

DEE H. ANDREWS
Technical Director

LYNN A. CARROLL, Colonel, USAF
Chief, Warfighter Training Research Division

Please notify AFRL/HEOP, 2509 Kennedy Circle, Building 125, Rm 158, Brooks AFB TX 78235-5118, if your address changes, or if you no longer want to receive our technical reports. You may write or call the STINFO Office at DSN 240-3877, or Commercial (210) 536-3877. E-mail address is shirley.walker@platinum.brooks.af.mil.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1998	Interim - July 1996 to September 1996

4. TITLE AND SUBTITLE

Near Identity of Cognitive Structure in Sex and Ethnic Groups

5. FUNDING NUMBERS

PE - 62205F
PR - 1123
TA - B1
WU - 01

6. AUTHOR(S)

Thomas R. Carretta
Malcolm J. Ree

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division, Training Effectiveness Branch
7909 Lindbergh Drive
Brooks AFB TX 78235-5352

8. PERFORMING ORGANIZATION

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6001 South Power Road, Bldg 558
Mesa AZ 85206-0904

10. SPONSORING/MONITORING

AL/HR-TP-1997-0006

11. SUPPLEMENTARY NOTES

Air Force Research Laboratory Technical Monitor: Dr Thomas R. Carretta, (510) 536-3622

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Comparisons of aptitude factor structures were made in large samples of young Americans who took a multiple aptitude test battery. The factor model that was used had been statistically confirmed. It included hierarchical g and five lower-order factors representing Verbal, Math, Spatial, Technical Knowledge, and Perceptual Speed. The model showed good fit for both the sexes (male and female) and ethnic groupWhite, Black, Hispanic, Asian American, and Native American) comparisons. The proportions of total and common variance accounted for by g and the five lower-order factors were similar for men and women and for all five ethnic groups. Confirmatory factor techniques that imposed statistical constraints tested if the loadings of the tests were the same for both sexes and for Whites versus each of the other ethnic groups. Although many of the tests of the differences in the loadings were statistically significant, most differences were small in magnitude (less than .05). The most notable differences occurred for a test of aviation knowledge. It had a lower loading for men than for women on both hierarchical g and on the lower-order Technical Knowledge factor. It also had a lower loading for Whites than for Blacks and Asian Americans on the lower-order Technical Knowledge factor. Correlations between factor loadings for the sex groups and for all pairs of ethnic groups were very high, approaching r = 1.0. Regressions between pairs of groups indicated that there was no mean difference in loadings between males and females and among the ethnic groups. These findings, along with previous research, present a consistent picture of near identity of structure of intellect for sex and ethnic groups.

14. SUBJECT TERMS

Academic aptitude; AFOQT; Air Force Officer Qualifying Test: Armed Services Vocational Aptitude Battery; ASVAB; Aviation knowledge; Ethnic groups; Intellect; Math ability; Perceptual speed; Sexes; Spatial ability; Technical knowledge; Verbal ability

15. NUMBER OF PAGES
19

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

CONTENTS

FIGURES

TABLES

PREFACE

This effort was conducted under Work Unit 1123-B1-01, Pilot Selection and Classification Support, which is dedicated to research into the selection and classification of U. S. Air Force aircrew personnel.

We would like to thank Paul J. Rioux for his technical assistance in preparing the data files, and Dr Joseph L. Weeks for his comments. This paper has been published previously as Carretta, T. R., & Ree, M. J. (1995), Near identity of cognitive structure in sex and ethnic groups. *Personality and Individual Differences*, *19*, 149-155.

Send written correspondence to AFRL/HEAB, 7909 Lindbergh Drive, Brooks AFB, TX 78235-5352. Send electronic mail to Thomas.Carretta@williams.af.mil.

NEAR IDENTITY OF COGNITIVE STRUCTURE IN SEX AND ETHNIC GROUPS

SUMMARY

Comparisons of aptitude factor structures were made in large samples of young Americans who took a multiple aptitude test battery. The factor model that was used had been statistically confirmed. It included hierarchical *g* and five lower-order factors representing Verbal, Math, Spatial, Technical Knowledge, and Perceptual Speed. The model showed good fit for both the sexes (male and female) and ethnic group (White, Black, Hispanic, Asian-American, and Native-American) comparisons. The proportions of total and common variance accounted for by *g* and the five lower-order factors were similar for men and women and for all five ethnic groups. Confirmatory factor techniques that imposed statistical constraints tested if the loadings of the tests were the same for both sexes and for Whites versus each of the other ethnic groups. Although many of the tests of the differences in the loadings were statistically significant, most differences were small in magnitude (less than .05). The most notable differences occurred for a test of aviation knowledge. It had a lower loading for men than for women on both hierarchical *g* and on the lower-order Technical Knowledge factor. It also had a lower loading for Whites than for Blacks and Asian-Americans on the lower-order Technical Knowledge factor. Correlations between factor loadings for the sex groups and for all pairs of ethnic groups were very high, approaching $\underline{r} = 1.0$. Regressions between pairs of groups indicated that there was no mean difference in loadings between males and females and among the ethnic groups. These findings, along with previous research, present a consistent picture of near identity of structure of intellect for sex and ethnic groups.

INTRODUCTION

There have been numerous studies of group differences in mean aptitude scores. Willerman (1979) reviewed several studies comparing average differences in factors for American and foreign racial and ethnic groups. Similarly, Loehlin, Lindzey, and Spuehler (1975) provided a summary of mean difference studies among groups. While informative, these studies did not address differences in factor structure among groups. The current effort remedies that deficit by providing an investigation of the structure of cognitive abilities for sex and racial/ethnic groups.

Studies of Mean Differences

There have been several studies of aptitude differences between sex and ethnic groups. For example, Hyde (1981) in a meta-analysis of male-female differences in cognitive abilities, reported median differences of .45s for verbal, .43s for quantitative, and .45s for visual-spatial ability. Lesser, Fifer, and Clark (1965) investigated differences among ethnic groups on abilities. They compared mean profiles of scores for Americans identified as Black, Chinese, Jewish, and

Puerto Rican. They found not only mean differences, but also differences in the pattern of mean profiles among the groups. For example, the Black and Jewish groups had their highest average scores on the verbal factor, while the Chinese and Puerto Rican groups had their lowest average scores on this factor. Space perception, in contrast, was relatively lower for the Black and Jewish groups and highest for the Chinese and Puerto Rican groups.

More recently, Lynn and Owen (1994) studied three ethnic groups in South Africa (Whites, Indians, and Blacks). The mean standardized White-Black difference was about 2.1 σ and the mean standardized White-Indian difference was 1.0 σ. In each case the White mean was higher. It should be noted that lack of English as a first language for the Black group may have contributed to the observed White-Black difference on the tests.

In American samples, similar patterns of intercorrelations among aptitude measures have been found for Blacks and Whites in studies by Nichols (1970) and Scarr-Salapatek (1971). Both studies concluded that it was unlikely that underlying dimensions of ability varied. These differences in dimensionality would give rise to different factor structures that could be statistically tested.

Studies of Factorial Similarity

Several studies of factorial similarity have been conducted. Michael (1949) compared the factor structure of World War II U. S. Army Air Force pilot candidate selection tests for Blacks and Whites and found virtually no differences. Humphreys and Taber (1973) compared factor structures for high and low socio-economic status boys from Project Talent and found no discernible differences. Loehlin, Lindzey, and Spuehler (1975) noted that although the ethnicity of the participants in Project Talent was not specifically identified, it could be expected that the ethnic composition of the two groups would differ significantly.

DeFries et al. (1974) compared the structure of ability for two ethnic groups as measured by 15 cognitive tests. Their participants were Hawaiians of either Japanese or European ancestry. They concluded that the same four factors were present in both groups and that the factor loadings were nearly identical for the two groups.

These previous studies all investigated the relationships among groups for common factors. Each of these common factors was mostly general cognitive ability (Jensen, 1980; Ree & Earles, 1991). More informative analyses were conducted by Ree and Carretta (1995) who investigated the comparative structure of another multiple aptitude battery, the ASVAB, across sex and ethnic groups using a hierarchical model of ability (see Vernon, 1969) and conducted statistical tests of the similarity of the factor loadings. They found only small differences on the Verbal/Math and Speed factors. However, there were notable differences on the Technical Knowledge factor. Most of the differences in this factor were due to a single specific knowledge test of electronics information.

The ASVAB is relatively factor poor. The Air Force Officer Qualifying Test (AFOQT) has a greater number of tests and a greater diversity of factors. Skinner and Ree (1987) conducted exploratory factor analyses of the AFOQT using a sample of 3,000 US Air Force officer applicants. The correlations among the 16 tests ranged from .17 to .77, with an average correlation of .43. Based on a principal factors analysis with communalities in the principal diagonal and Kaiser-Harris Type II oblique rotation, they described a 5-factor solution: Verbal, Math, Spatial, Technical Knowledge (Aircrew Interest/Aptitude), and Perceptual Speed. These factors showed an average correlation of .36 with a range of .22 to .50. Although the sample contained both men and women and several ethnic groups, no group-specific analyses were conducted.

Earles and Ree (1991) in a reanalysis of the Skinner and Ree (1987) data, demonstrated the hierarchical structure of the AFOQT. They found an hierarchical estimate of general cognitive ability, *g*, based on correlated lower-order factors. As with Skinner and Ree (1987), they did not test the fit of the model or perform group specific analyses.

More recently, Carretta and Ree (1996) used confirmatory factor methods (Bentler, 1989) to test the structure of the AFOQT in the Skinner and Ree (1987) sample. They examined seven structural models including: *g* only, two models based on the operational AFOQT composites (not including the Academic Aptitude composite which would have been redundant with the Verbal and Quantitative composites and would have caused a linear dependency in the data), two models based on Skinner and Ree (1987), and two models that simplified the Skinner and Ree (1987) solution by combining some of the lower-order factors.

The two models based on operational composites were a 4-factor solution (Verbal, Quantitative, Pilot, and Navigator-Technical) and the same with a hierarchical *g* added. The two Skinner and Ree (1987) models were 5-factors (Verbal, Math, Spatial, Technical Knowledge, and Perceptual Speed) and 5-factors with a hierarchical *g* added. Model 6 was a simplification of Skinner and Ree (1987) with hierarchical *g* that combined the Spatial and Perceptual Speed factors, but kept all the others. Model 7 combined the Verbal and Math factors to produce a Vernon-like model (Vernon, 1969). This resulted in a model with hierarchical *g* and residualized lower-order Verbal/Math (combined from Skinner & Ree, 1987), Spatial/Perceptual Speed (combined from Skinner & Ree, 1987), and Technical Knowledge factors. The Skinner and Ree (1987) model with hierarchical *g* and five 5 lower-order factors was selected as it showed both good fit and interpretability. The Carretta and Ree (1996) analyses did not address possible group differences as sex and ethnic group comparisons were not made.

The current study investigated whether the factor structure of abilities was the same for men and women and for five ethnic groups. If the factors or loadings of tests on the factors differ among or between groups, scores based on the factors cannot have the same meaning in research or in practical application. Differences in factor structure or loadings could occur as a result of differences in personal choice (e.g., course enrollment, leisure interests) and opportunity.

The consequences of group differences in aptitude measurement on operational scores are a function of the magnitude of the differences in factor loadings and the number of factors that contribute to operational scores. These consequences extend to using the test for research purposes. Finally, differences in factor loadings cannot necessarily be interpreted to indicate test bias.

The Carretta and Ree (1996) factor structure was used as the basis for the current study as it had been statistically confirmed and was meaningfully interpretable. These qualities made it superior to the other models. The goal of this study was to determine if the factor structure was the same for men and women and for five ethnic groups.

METHOD

Participants

The participants were 269,968 applicants for US Air Force officer commissions tested between 1981 and 1993. They were 81.4% male and 78.7% White, 12.2% Black, 4.7% Hispanic, 3.5% Asian-American, and 0.9% Native-American. Their ages at time of testing ranged between 18 and 27 years. The age, sex, and ethnicity of the participants were collected by self-report.

Measures

The AFOQT is a multiple-aptitude test battery used for the commissioning of US Air Force officers through Officer Training School (OTS) and the Reserve Officer Training Corps (ROTC). AFOQT scores are used along with other measures of aptitude and educational achievement (e.g., college grade point average, type of degree, previous flying experience) to qualify applicants who pass medical and physical requirements for pilot and navigator training. The AFOQT is similar to many other multiple aptitude tests such as the Differential Aptitude Tests, General Aptitude Test Battery, and the Armed Services Vocational Aptitude Battery because it contains measures of verbal, mathematical, and spatial factors. It differs from the others in its measurement of specialized aviation knowledge. All of these tests stem from the theoretical model offered by Kelley (1928) and contain several of the factors he posited such as spatial, quantitative, verbal, and speed.

The AFOQT consists of 16 tests that include Verbal Analogies (VA), Arithmetic Reasoning (AR), Reading Comprehension (RC), Data Interpretation (DI), Word Knowledge (WK), Math Knowledge (MK), Mechanical Comprehension (MC), Electrical Maze (EM), Scale Reading (SR), Instrument Comprehension (IC), Block Counting (BC), Table Reading (TR), Aviation Information (AI), Rotated Blocks (RB), General Science (GS), and Hidden Figures (HF). A more detailed description of the tests was provided elsewhere (Carretta & Ree, 1995; Skinner & Ree, 1987). Its reliability has been examined (Earles & Ree, 1991; Skinner & Ree, 1987) and it has been validated for pilot and navigator training (Arth, Steuck, Sorrentino, & Burke, 1990;

Carretta, 1992; Carretta & Ree, 1994, 1995, Koonce, 1982; Olea & Ree, 1994). All 16 AFOQT tests were used in this study.

Procedures

The factor structure tested across groups in this study was confirmed by Carretta and Ree (1996) and represented a 6-factor solution (See Figure 1). The hierarchical factor is *g* (all 16 tests) and the five lower-order factors are Verbal (VA, RC, WK, GS), Math (AR, DI, MK, SR), Spatial (MC, EM, BC, RB, HF), Technical Knowledge (MC, IC, AI, GS), and Perceptual Speed (DI, SR, BC, TR). The model is specified in residualized form (Schmid & Leiman, 1957) with the effects of the hierarchical factor removed from the five lower-order factors. Maximum likelihood estimation procedures were used.

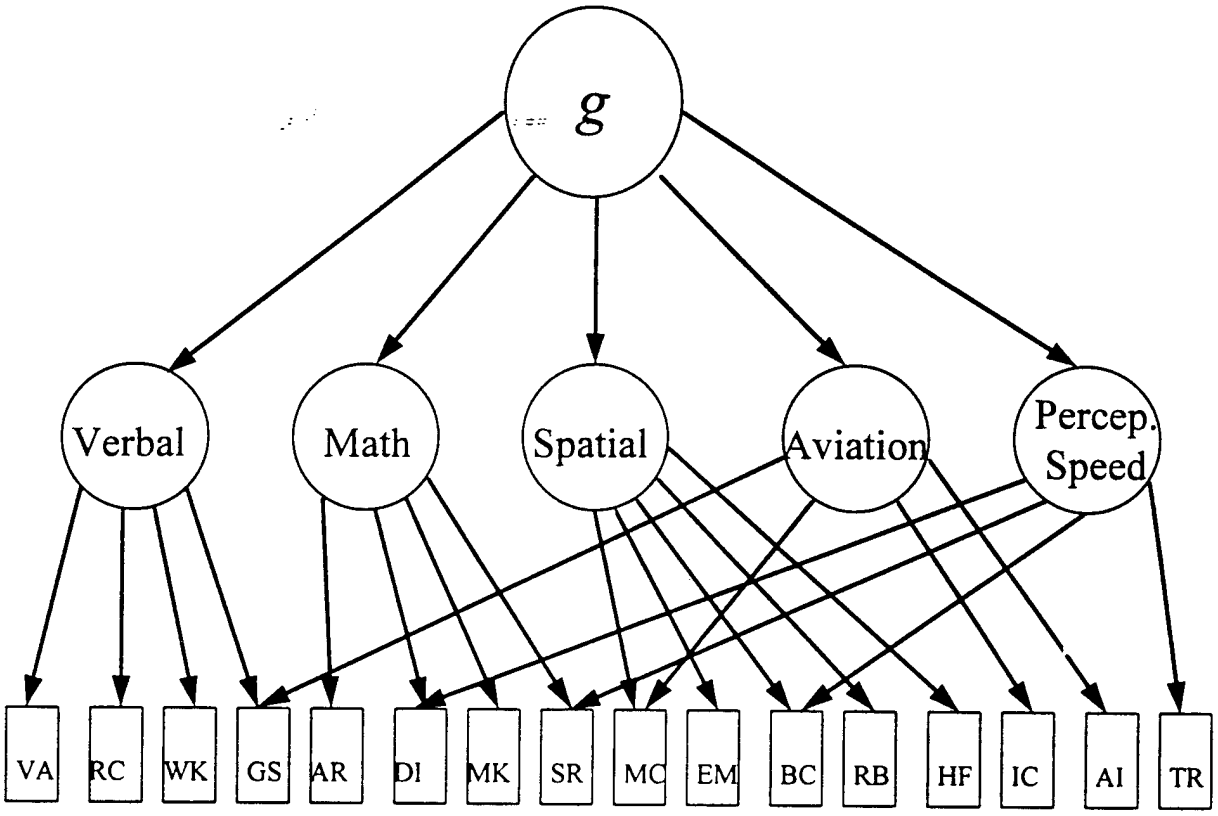


Figure 1. Structural Model of the Air Force Officer Qualifying Test

A chi-square test was conducted to determine if the loadings for a variable on a factor for two groups were the same (Bentler, 1989). To determine if the factor loadings for a test were the same for two groups (e.g., Whites and Blacks), it is necessary to estimate the structural model parameters with the loadings for both groups constrained to the same value and then to estimate the model parameters with the constraint removed. This procedure lets each group-specific (e.g.,

Whites or Blacks) factor loading estimate have a different value. If the chi-square test is significant, separate parameter estimates for each group provide a better fit and are statistically justifiable. In this study, the loadings of males versus females for each test on each of the six factors were tested. The most important ethnic comparisons are those between the majority and each minority group. The tests of equality of the loadings were done by separately comparing the majority group (Whites) with each of the four minority groups: Whites versus Blacks, Whites versus Hispanics, Whites versus Asian-Americans, and Whites versus Native-Americans. A Type I error rate of $p < .01$ was used for all statistical tests.

Several goodness-of-fit statistics were considered. These included the Bentler-Bonett Non-Normed Fit Index (BBNNI) (Bentler, 1989), the Tucker-Lewis Incremental Fit Index (TLI) (Tucker & Lewis, 1973), the root mean square error of approximation (RMSEA) (Browne & Cudeck, 1993), and the Comparative Fit Index (CFI) (Bentler, 1990). The BBNNI has been shown to be susceptible to sample size effects (Marsh, Balla, & McDonald, 1988). Bentler (1989) extended the TLI to develop the CFI, which has been shown to be less susceptible to the effects of sample size with a smaller sampling variance than the TLI. As a result, the CFI was chosen to evaluate the goodness-of-fit of the models. Additionally, the RMSEA that gives an indication of amount of error per degree of freedom, is presented. The models were evaluated with all constraints in place.

Pearson correlations of factor loadings for pairs of groups (males and females, all pairs of ethnic groups) were examined to further explore the relationship of factor loadings between groups. Regressions were computed and the regression constant (y intercept) was evaluated to determine whether a constant bias existed.

RESULTS

No special problems occurred in estimating the parameters for the models. The CFI value for fitting the male versus female comparative model was .936 with a RMSEA of .052 (χ^2 = 143,522, df = 203). The CFI value for fitting the 5 group ethnicity comparison was .927 with a RMSEA of .031 (χ^2 = 146,089, df = 563).

The percent of total variance accounted for by each factor is shown for each group in Table 1. The proportion of total variance accounted for by *g* was similar for males (38%) and females (40%). The proportion of common variance accounted for by *g* also was similar for males and females (66% versus 67%). The proportions of total variance accounted for by *g* were 33% for Whites, 35% for Blacks, 34% for Hispanics, 33% for Asian-Americans, and 33% for Native-Americans. The proportions of common variance for which *g* accounted for these groups were similar: 60%, 59%, 59%, 59%, and 59%. The proportions of total and common variance accounted for by the five lower-order factors also were similar for both sexes and across all ethnic groups.

Table 1.
Percent of Total Variance Accounted for by Factors Within the Groups

Group	Percent of Total Variance					
	g	Verbal	Math	Spatial	Technical Knowledge	Perceptual Speed
Males	38.15	6.81	3.59	2.58	4.30	2.69
Females	39.97	5.16	3.62	2.75	5.57	2.58
Whites	33.05	7.81	3.78	2.90	5.11	2.88
Blacks	35.42	7.94	3.92	3.20	6.41	2.79
Hispanics	33.52	7.94	3.88	2.92	5.42	2.78
Asian-Americans	32.93	7.79	3.73	2.87	5.90	2.75
Native-Americans	33.00	7.88	3.73	2.87	5.31	2.77

The loadings of the tests on the factors for males and females are shown in Tables 2 (*g*) and 3 (five lower-order factors). Although many of the univariate tests for differences in factor loadings between males and females were statistically significant (30 of 37), only four pairs of loadings differed by more than .05. These were the loadings for EM, IC, and AI on *g* and for AI on the lower-order Technical Knowledge factor. AI showed the largest male/female difference, .684 versus .800 on this lower-order factor.

Tables 2 (*g*) and 3 (five lower-order factors) also show the loadings for the five ethnic groups. Although many of the univariate tests for differences in loadings between Whites and the other ethnic groups were statistically significant, few were larger than .05 in magnitude. These included EM and IC on *g* for Whites versus Blacks, EM on the lower-order Spatial factor for Whites versus Blacks, and AI on the lower-order Technical Knowledge factor for Whites versus Blacks and Whites versus Asian-Americans. The largest differences in loadings occurred for the AI test on the lower-order Technical Knowledge factor (Whites, .695; Blacks, .795; Hispanics, .717; Asian-Americans, .772; Native-Americans, .709).

Pearson correlations between factor loadings for the sex groups and among the ethnic groups were very high, approaching $\underline{r} = 1.0$. The correlation of the loadings for males and females for the hierarchical *g* factor was .97 and ranged from .97 to .99 for pairs of the five ethnic groups. Results for the residualized lower-order factors showed the same range.

Table 2.
Factor loadings by Sex and Ethnicity for the Hierarchical *g*

	Sex		Ethnicity				
Test	Male	Female	White	Black	Hispanic	Asian-Am.	Native-Am.
VA	.715	.712	.635	.618	.623	.619	.627
AR	.753	.759	.730	.738	.738	.725	.724
RC	.615	.619	.521	.520	.513	.521	.516
DI	.710	.705	.667	.692	.682	.667	.673
WK	.518	.512	.432	.442	.446	.435	.441
MK	.723	.724	.707	.730	.710	.705	.705
MC	.606	.637	.556	.584	.568	.564	.564
EM	.457	.539	.448	.524	.477	.457	.456
SR	.737	.744	.706	.730	.713	.693	.705
IC	.604	.661	.572	.646	.589	.585	.585
BC	.660	.652	.615	.633	.607	.609	.610
TR	.534	.512	.463	.432	.446	.446	.446
AI	.375	.439	.326	.372	.336	.362	.332
RB	.611	.634	.588	.601	.574	.577	.578
GS	.596	.628	.570	.612	.588	.583	.580
HF	.527	.531	.494	.485	.488	.484	.484

Regressions were computed between the sex groups and between all pairs of the ethnic groups. All the regressions yielded intercepts of approximately zero, indicating that there was no mean difference in loadings.

DISCUSSION

Fit statistics indicated that the data gave a good fit to the model for both sexes and for the five ethnic groups. The AFOQT is somewhat less *g*-saturated than the military enlistment test battery, the Armed Services Vocational Aptitude Battery (Carretta & Ree, 1996). Hierarchical *g* accounted for the greatest amount of variance for both sexes and all ethnic groups and differed little by sex or ethnicity. Further, *g* accounted for more of the total variance than did the sum of the five lower-order factors. This is consistent with results for other aptitude tests (Jensen, 1980; Ree & Carretta, 1994, 1995).

Table 3.
Factor Loadings by Sex and Ethnicity for the Residualized Lower-Order Factors

	Sex		Ethnicity				
Factor/Test	Male	Female	White	Black	Hispanic	Asian-Am	Native-Am
Verbal							
VA	.406	.404	.464	.451	.455	.452	.458
RC	.573	.577	.630	.628	.621	.629	.624
WK	.736	.728	.764	.781	.788	.768	.778
GS	.234	.247	.232	.249	.239	.237	.236
Math							
AR	.646	.651	.667	.675	.675	.662	.661
DI	.298	.296	.310	.321	.317	.310	.312
MK	.214	.214	.206	.213	.207	.206	.205
SR	.148	.149	.146	.151	.147	.143	.145
Spatial							
MC	.197	.207	.241	.253	.246	.244	.244
EM	.265	.313	.292	.342	.312	.299	.298
BC	.278	.274	.281	.289	.277	.278	.279
RB	.391	.405	.422	.432	.413	.415	.415
HF	.270	.272	.253	.248	.250	.247	.248
Technical Knowledge							
MC	.301	.317	.371	.390	.379	.376	.376
IC	.284	.311	.337	.381	.347	.345	.345
AI	.684	.800	.695	.795	.717	.772	.709
GS	.222	.234	.288	.310	.298	.295	.294
Perceptual Speed							
DI	.094	.093	.106	.110	.109	.106	.107
SR	.323	.326	.322	.333	.326	.317	.322
BC	.328	.324	.338	.348	.334	.335	.335
TR	.458	.440	.481	.449	.463	.464	.464

Among the residualized lower-order factors, the following stand out. The two largest sources of test variance for all groups are Verbal and Technical Knowledge. The Verbal factor accounted for more variance than did the Technical Knowledge factor for all groups except females, where the reverse was true. However, the most notable finding was the great similarity in percent-of-total-variance-accounted-for by each of the factors for all groups.

The similarity in the proportions of total and common variance for *g* for the sex and ethnic groups suggested that the combination of unreliability and unique variance was similar across groups. Also, the proportions of variance accounted for by the lower-order factors were similar across groups. The largest sex and ethnic group differences occurred for the Technical Knowledge factor which accounted for less total and common-variance for males (4.3% and 7.4%) than for females (5.6% and 9.3%) and less for Whites (5.1% and 9.2%) than for Blacks (6.4% and 10.7%), Hispanics (5.4% and 9.6%), Asian-Americans (5.9% and 10.5%), and Native-Americans (5.3% and 9.6%). It is likely that these differences can be attributed to opportunity or individual decisions to learn aviation principles and flying.

The most notable differences in factor loadings occurred for the Electrical Maze (EM), Instrument Comprehension (IC), and Aviation Information (AI) tests. All three of these tests had lower loadings for males than for females on *g* and AI had a lower loading for males on the lower-order Technical Knowledge factor. These differences in loadings are most likely the consequences of personal choice and opportunity. For example, Ree and Carretta, 1994) found differences in loadings for males and females on technical knowledge tests on a similar battery and attributed them to differences in high school course enrollment patterns.

For the ethnic group comparisons, EM and IC had lower loadings on *g* for Whites than for Blacks. EM also had a lower loading on the lower-order Spatial factor for Whites than for Blacks. On the lower-order Technical Knowledge factor, AI had a lower loading for Whites than for Blacks and Asian-Americans. The reasons for these differences are not apparent. The correlational and regression analyses were consistent with the findings of the comparisons of the structural model.

The results of this study add to the findings (DeFries et al., 1974; Michael, 1949) that the structure of cognitive ability is nearly identical across sex and ethnic groups. These results further expand our understanding by including additional ethnic groups and by using a hierarchical model of human abilities.

REFERENCES

Arth, T. O., Steuck, K. W., Sorrentino, C. T., & Burke, E. F. (1990). *Air Force Officer Qualifying Test (AFOQT): Predictors of undergraduate pilot training and undergraduate navigator training success* (AFHRL-TP-89-52). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Bentler, P. M. (1989). *EQS structural equations program manual*. Los Angeles, CA, BMDP Statistical Software.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen and J. S. Lang (Eds.), *Testing structural equation models*. Newbury Park, CA: Sage.

Carretta, T. R. (1992). Understanding the relations between selection factors and pilot training performance: Does the criterion make a difference? *The International Journal of Aviation Psychology, 2*, 95-105.

Carretta, T. R., & Ree, M. J. (1994). Pilot-candidate selection method: Sources of validity. *The International Journal of Aviation Psychology, 4*, 103-117.

Carretta, T. R., & Ree, M. J. (1995). Air Force Officer Qualifying Test validity for predicting pilot training performance. *Journal of Business and Psychology, 9*, 379-388.

Carretta, T. R., & Ree, M. J. (1996). Factor structure of the Air Force Officer Qualifying Test: Analysis and comparison. *Military Psychology, 8*, 29-42.

DeFries, J. C., Vandenberg, S. G., McClearn, G. E., Kuse, A. R., Wilson, J. R., Ashton, G. C., & Johnson, R. C. (1974). Near identity of cognitive structure in two ethnic groups. *Science, 183*, 338-339.

Earles, J. A., & Ree, M. J. (1991). *Air Force Officer Qualifying Test (AFOQT): Estimating the general ability component* (AL-TP-1991-0039). Brooks AFB, TX: Manpower and Personnel Research Division, Human Resources Directorate, Armstrong Laboratory.

Humphreys, L. G., & Taber, T. (1973). Ability factors as a function of advantaged and disadvantaged groups. *Journal of Educational Measurement, 10*, 107-115.

Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using ω^2 and *d*. *American Psychologist, 36*, 892-901.

Jensen, A. R. (1980). *Bias in mental testing*. New York: The Free Press.

Kelley, T. L. (1928). *Crossroads in the mind of man: A study of differentiable mental abilities*. Palo Alto, CA: Stanford University Press.

Koonce, J. M.. (1982). Validation of a proposed pilot-trainee selection system. *Aviation, Space, and Environmental Medicine, 53*, 1166-1169.

Lesser, G. S., Fifer, G., & Clark, D. H. (1965). Mental abilities of children from different social-class and cultural groups. *Monographs of the Society for Research in Child Development, 30*.

Loehlin, J. C., Lindzey, G., & Spuehler, J. N. (1975). *Race differences in intelligence*. San Francisco: Freeman.

Lynn, R., & Owen, K. (1994). Spearman's hypothesis and test score differences between Whites, Indians, and Blacks in South Africa. *The Journal of General Psychology, 121*, 27-36.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391-410.

Michael, W. B. (1949). Factor analyses of tests and criteria: A comparative study of two AAF pilot populations. *Psychological Monographs, 63*, 55-84.

Nichols, P. L. (1970). The effects of heredity and environment on intelligence test performance in 4 and 7 year white and Negro sibling pairs. Unpublished doctoral dissertation. University of Minnesota, Ann Arbor, MI: University Microfilms 71-18,874.

Olea, M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than *g*. *Journal of Applied Psychology, 79*, 845-851.

Ree, M. J., & Carretta, T. R. (1994). Factor analysis of the ASVAB: Confirming a Vernon-like structure. *Educational and Psychological Measurement, 54*, 459-463.

Ree, M. J., & Carretta, T. R. (1995). Group differences in aptitude factor structure. *Educational and Psychological Measurement*, 55, 268-277.

Ree, M. J., & Earles, J. A. (1991). Predicting training success: Not much more than *g*. *Personnel Psychology, 44*, 321-332.

Scarr-Salapatek, S. (1971). Race, social class, and IQ. *Science, 174*, 1285-1295.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53-61.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analysis of form O* (AFHRL-TR-86-68). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10.

Vernon, P. E. (1969). *Intelligence and cultural environment*. London: Methuen.

Willerman, L. (1979). *The psychology of individual and group differences.* San Francisco CA: W.H. Freeman